

09/882,563

Remarks

Reconsideration of the above referenced application in view of the enclosed amendment and remarks is requested. Claims 1, 3-5, 13, 18, 23, and 27 have been amended. Existing claims 1 to 30 remain in the application. Claims 31-46 have been added to claim additional aspects of the invention as described in the specification as originally filed.

The title has been amended at the Examiner's request to be more descriptive of the claimed invention. The specification has been amended to correct typographical errors, as suggested by the Examiner. The specification has been amended to clearly reference items 20 and 38. These elements are described in the specification as originally filed as well as shown in the Figures. No new matter is introduced.

Figure 2 is to be replaced with Figures 2A and 2B, at the Examiner's request. The replacement Figures are attached, as well as a copy of the original Figure 2. The specification is also amended to reference Figures 2A and 2B, where appropriate, instead of referencing deleted Figure 2. Applicants assert that it is standard practice to include optional process blocks within the same flow chart as mandatory processes. It is understood by one of ordinary skill in the art that various representations may be used to indicate that some processes are optional. One method is to surround the process block with a dotted or dashed line instead of a solid line, as now the case for block 46. Another method is to have directional lines leading in two directions to two processes (as was the case in Figure 2 coming from 32). The specification, as originally filed, clearly describes the two embodiments as shown in Figure 2 and now shown in Figures 2A and 2B. The drawings have been modified to make the distinction between the two embodiments more clear to the Examiner and do not introduce new matter, nor does this modification change the scope of the claimed inventions.

09/882,563

Applicants wish to thank the Examiner for the indication that Claim 4 is directed to allowable subject matter. Claim 4 has been amended to put it into independent form, and as such, should be allowed to issue at the earliest possible time.

ARGUMENT

The amendments to the drawings and specification, as described above make Examiner's objections moot.

Claims 5 and 10 (and 16) are objected to by the Examiner. Claim 5 has been amended, as suggested to correct a typographical error. This amendment neither broadens nor narrows the scope of Claim 5 and its equivalents. The Examiner's objection to Claim 10 is respectfully traversed and Claim 10 is believed allowable based on the above amendments and following discussion.

The Examiner asserts that the term "recognizer information" has no clear interpretation. Applicants assert that the term "recognizer information" is clearly defined in the specification as originally filed, and that the Examiner's assumptions are incorrect. Specifically, on page 9, recognizer information is described as: "...associated recognizer information that may include individual result confidence values...." On page 12, Figure 2 (original), recognizer information is further described. Specifically block 46, "Update Recognizer Information" is described in the context of one embodiment of the invention. Recognizer information is described as information about the recognizer and not merely information from the recognizer as assumed by the Examiner. Specifically, the "feedback" (44) is described as "information regarding the correctness of the recognition process," (page 7) and is necessary to tune the predictor over time. This feedback may then be "used to update information about the recognizers at 46, such as their performance with a given characteristics in the contextual information, a weighting that is applied to their individual-result confidence values, their optimization for various channels and devices, among many others." (page 12). As described by the specification as originally filed, "recognizer information" may be information about the environment of the recognizer, meta-information

09/882,563

about results of the recognizer such as confidence values, or other related information.
[Emphasis added]

Claims 1-3, 5-7, 9-10, 13-16, 18-19, 21, 23-24 & 27-29 are rejected under 35 U.S.C. §103(a) as being unpatentable over Kenne et al. ("Hybrid Language Models and Spontaneous Legal Discourse", 4th International Conference on Spoken Language, October 1996) (hereafter, Kenne et al.) in view of Hsu et al. (US 5,677,991 A) (hereafter Hsu et al.). This rejection is respectfully traversed and Claims 1-3, 5-7, 9-10, 13-16, 18-19, 21, 23-24 & 27-29 are believed allowable as amended based on the foregoing and following discussion.

Regarding, Claims 1, 13, 18, 23 and 27, the Examiner asserts that Kenne et al. teaches selecting a recognizer; receiving an input stream; deriving selection information, wherein the selection information includes performance-related information; using the selection information to select results from at least one enabled recognizer. The Examiner admits that Kenne et al. does not mention applications, but asserts that this element is taught by Hsu et al.

Claims 1, 13, and 18 are amended to require that a recognizer is enabled based upon an expected future performance of the recognizer. This Examiner has indicated that this limitation is directed toward allowable subject matter. Therefore Claims 1, 13, 18 and their progeny are allowable. Applicants respectfully request that the Examiner allow these claims to issue at the earliest possible time.

Claims 23 and 27 require using the enabling information to select an enabled recognizer. This limitation is neither taught nor suggested by Kenne et al. or Hsu et al. Kenne et al. does not teach to enable recognizers. The models used in Kenne et al. are turned on one at a time. The local perplexity determines the selection of the model (or recognizer). Kenne et al. teaches the training of three models, Both, L+W, and Hybrid. In operation, the Kenne et al. speech recognizer takes an input stream and applies it to one model, based on the source (lawyer or witness). In contrast, Applicants claimed invention selects which recognizers to send the input stream to, either in parallel or in sequence, using enabling information. More than one recognizer may be enabled. The results are selected from one of the enabled recognizers based on selection information in the predictor. By selectively enabling recognizers, one or more recognizers may be omitted from the selection of the

09/882,563

results. This is important to override a recognizer when it may otherwise be the one selected to have the best results. Enabling information, as defined by Applicants is at least described on page 4 of the specification. It is described that:

“For example, at 32, the input stream may be analyzed to determine characteristics of the communication channel. In some systems, analyzers exist that allow the system to determine the audio characteristics of the channel, for example, determining if the cellular or landline communication networks are in use. Other information may also be derived, including background noise and signal strength among others. Additionally, this analysis may determine characteristics of the communication device, for example determining if a speaker-phone or a wireless handset are in use. This analysis may occur as part of deriving enabling information for the recognizers.” [emphasis added]

The effect of disabling certain recognizers, i.e., those not enabled, is not the same as selecting one specific model (recognizer) as is taught by Kenne et al. Kenne et al. has no mechanism for taking results from more than one recognizer at the same time and predicting the best results. Therefore, Claims 23 and 27 are believed allowable, as amended, and should be allowed to issue.

Regarding Claim 2, the Examiner asserts that Kenne et al. teaches the feature that causes a recognizer to be selected that is different from the recognizer used in a previous interaction. The Examiner admits that Kenne et al. does not teach that the selection information is updated, but asserts that this is taught by Hsu et al. by setting a reference score as a baseline that is necessary to provide further evaluation. This rejection is respectfully traversed based on the foregoing amendments and foregoing and following discussion. Claim 2 is allowable because it is dependent on a claim with allowable subject matter.

Regarding Claim 3, the Examiner asserts that Kenne et al. discloses the feature of deriving enabling information, and using the information to enable at least one selected recognizer to process the input stream. This rejection is respectfully traversed and Claim 3 is believed allowable as amended based on the foregoing and following discussion. Claim 3

09/882,563

requires deriving enabling information, and using the enabling information to enable at least one selected recognizer. This limitation is neither taught nor suggested by Kenne et al. or Hsu et al. Kenne et al. does not teach deriving enabling information, nor actually enabling at least one selected recognizer, as previously discussed with reference to claims 23 and 27. Therefore, Claim 3 is believed allowable as amended.

Regarding Claims 5, 6 and 24, the Examiner asserts that Results section of Kenne et al. discloses the feature that the enabling information, and consequently the performance-related information, comprises at least one type of information from the group comprised of: channel characteristics, user information, contextual information, dialog state, recognizer costs and performance history. This rejection is respectfully traversed based on the foregoing and following discussion. Neither Kenne et al., nor Hsu et al. teach deriving, analyzing or using enabling information, especially recognizer costs. Applicants describe that one element of performance-related information is a quantitative analysis of the costs of using a particular recognizer, in either financial or computational terms. This information is part the enabling information. Applicants use the enabling information to determine which recognizers to enable. Only the enabled recognizers process the input stream. The results of the enabled recognizers are then selected based on performance-related information. This is neither taught nor suggested by the cited references. Thus, claims 5, 6 and 24 are believed allowable.

Regarding Claims 7, 14 and 28, the Examiner asserts that Kenne et al. teaches deriving the selection information further comprises analyzing the input stream for channel characteristics. Claims 7 and 14 are dependent on a claim with allowable subject matter. Therefore, Claims 7 and 14 are allowable. Further, with respect to Claims 7, 17 and 28, Kenne et al. does not teach analyzing the input stream for channel characteristics as defined by Applicants. Kenne et al. merely determines on which track the transcript to be recognized is located and selects a model accordingly. Kenne et al does not "determine the audio characteristics of the channel, for example, determining if the cellular or landline communication networks are in use." Kenne et al. does not derive or analyze background noise and signal strength or other channel characteristics. The analysis as described in the specification may determine characteristics of the communication device, for example

09/882,563

determining if a speaker-phone or a wireless handset are in use. Network-based information services such as CallerID in conjunction with a local or network-based database mapping calling number to channel and device characteristics may be utilized for similar effect. The cited references do not teach or suggest analyzing the input stream for channel characteristics to derive selection information. Claim 28 is also allowable based on the foregoing discussion of claim 27 on which it is dependent.

Regarding claims 9, 15, 19 and 29, the Examiner asserts that Kenne et al. discloses receiving contextual information associated with the input stream by virtue of distinguishing between Lawyers and Witnesses. Claims 9, 15, and 19 are dependent on a claim with allowable subject matter. Therefore, Claims 9, 15, and 19 are allowable. Claim 29 is allowable based on the foregoing and following discussion. Kenne et al. merely discloses being able to distinguish between two tracks of information to determine which recognition model to use. Applicants' claimed invention requires that selection information comprises receiving contextual information associated with the input stream. Contextual information is defined in the specification as originally filed as that information related to the environment around the input stream, including characteristics of the user and information derived from the call using network services such as CallerID. This information may be obtained dynamically or may be predetermined. Contextual information may include gender, age, ethnicity, whether the speaker speaks the language of the recognizers as a first language, among other personal information about the user. Also, the channel and device characteristics may be included in the contextual information. Kenne et al. does not derive contextual information as defined by the Applicants.

Regarding Claims 10 and 16, the Examiner asserts that Kenne et al. teaches the feature of receiving recognizer information from the enabled recognizers to be used in the selection information with respect to the switching described in Section 4 (Results) of Kenne et al. Kenne et al. does not receive recognizer information from the enabled recognizers to be used in the selection information. First, Kenne et al. does not teach enabled recognizers, as defined by Applicants. Kenne et al. does not teach or disclose recognition information as defined by Applicants, and as discussed above. Thus, Kenne et al, does not teach or disclose

09/882,563

receiving recognizer information from the enabled recognizers to be used in selection information. Further, Claims 10 and 16 are allowable as being dependent upon a claim with allowable subject matter.

Regarding claim 21, the Examiner asserts that Kenne et al. teaches that the predictor is operable to select a recognizer based upon the converted stream. Claim 21 is allowable as being dependent upon a claim with allowable subject matter.

Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over Kenne et al. in view of Hsu et al. (US 5,677,991 A) and further in view of Waibel et al. (U.S. Patent 5,712,957 A) (hereafter Weibel et al. '957). The Examiner admits that neither Kenne et al. nor Hsu et al. teach separate input devices. The Examiner asserts that Waibel et al. ('957) discloses different inputs and uses information which corresponds to analyzing the input stream for device characteristics. This rejection is respectfully traversed and Claim 8 is believed allowable based on the foregoing and following discussion.

The Examiner references elements 23 and 24 in Figure 1 of Weibel et al. '957 to show multiple input streams. This reference is erroneous because element 24 is not an input stream that could be recognized by a speech recognizer. It is not speech or in the same category of input as the other input stream. In Col. 4, lines 53-60, Weibel et al. '957 describes 24 as a "touch sensitive pad" or other input transducer. Weibel et al. '957 describes translating this input with a handwriting recognition engine or other device entry. This input is to be used to assist with correction and repair of the module (recognizer). It is not an alternative input stream to be analyzed for characteristics that will derive selection information derived from the device characteristics. It is an additional input stream that is used to correct the first input stream with related information. Further, Claim 8 is allowable based on its dependence from a claim with allowable subject matter.

Claims 11-12, 17, 25-26 and 30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kenne et al. in view of Hsu et al. (US 5,677,991 A) and further in view of Kundu (U.S. Patent 5,924,066 A) (hereafter Kundu). This rejection is respectfully traversed based on the foregoing and following discussion.

09/882,563

Regarding Claims 11, 17, 25 and 30, the Examiner admits that Kenne et al. and Hsu et al. do not disclose feedback. The Examiner asserts that Kundu teaches this with classifying a speech signal that receiving feedback and including the feedback in the selection information is disclosed. Kundu does not teach or disclose feedback as described by Applicants. Kundu teaches using a neural net to train the recognizer with data. Kundu describes a 3-layer neural net as follows:

"The learning law for the perceptron 40 is a simple error feedback. The network learns the associations between input and output patterns by being exposed to many lessons. The weights are adjusted until the desired target output is produced. This weight adaptation is referred to as error backpropagation learning law."

Kundu discloses using an error backpropagation algorithm in the neural net. This does not generate "feedback." A neural net, as described by Kundu, cannot update recognizer information as defined in Applicants' claimed invention, but it merely provides a weighting for the input factors and provides a likely output. The output of the middle layer of the neural net, which is the heart of the error backpropagation algorithm, is not human readable/comprehensible. It does not provide feedback which can be used in the selection information. The top layer (output) of the neural net provides the best guess at the result (i.e. speech components) and does not provide information about performance or other recognizer information. Applicants' feedback provides input to a predictor, or selection information, which is used to select the recognizer from a plurality of recognizers operating on the same input stream, either in parallel, or in sequence. Kundu teaches a neural net which uses an error backpropagation method to train one recognizer. Thus, the error backpropagation does not provide feedback which will result in Applicants' claimed invention.

Regarding Claims 12 and 26, the Examiner asserts that Kundu teaches that feedback is received from one of the group comprised of: off-line analysis, user feedback, and feedback from the recognizer. The neural net as described by Kundu provides a "simple error feedback." Error feedback in the context of a neural net is error associated with the results as they correspond to the training data. When in recognition mode, rather than training mode,

09/882,563

the error feedback is already programmed into the neural net and feedback with respect to off-line analysis, user feedback, and feedback from the recognizer is not provided. Thus, claims 12 and 26 are believed allowable.

Claims 20 and 22 are rejected under 35 U.S.C.103(a) as being unpatentable over Kenne et al. in view of Hsu et al. (US 5,677,991 A) and further in view of Waibel et al. (U.S. Patent 5,855,000 A) (hereafter Waibel et al. '000). This rejection is respectfully traversed and Claims 20 and 22 are believed allowable based on the foregoing discussion as being dependent upon a claim with allowable subject matter.

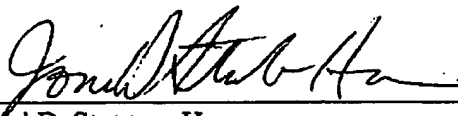
09/882,563

CONCLUSION

In view of the foregoing, claims 1-46 are all in condition for allowance. If the Examiner has any questions, the Examiner is invited to contact the undersigned at (703) 633-6845. Early issuance of Notice of Allowance is respectfully requested.

Respectfully submitted,

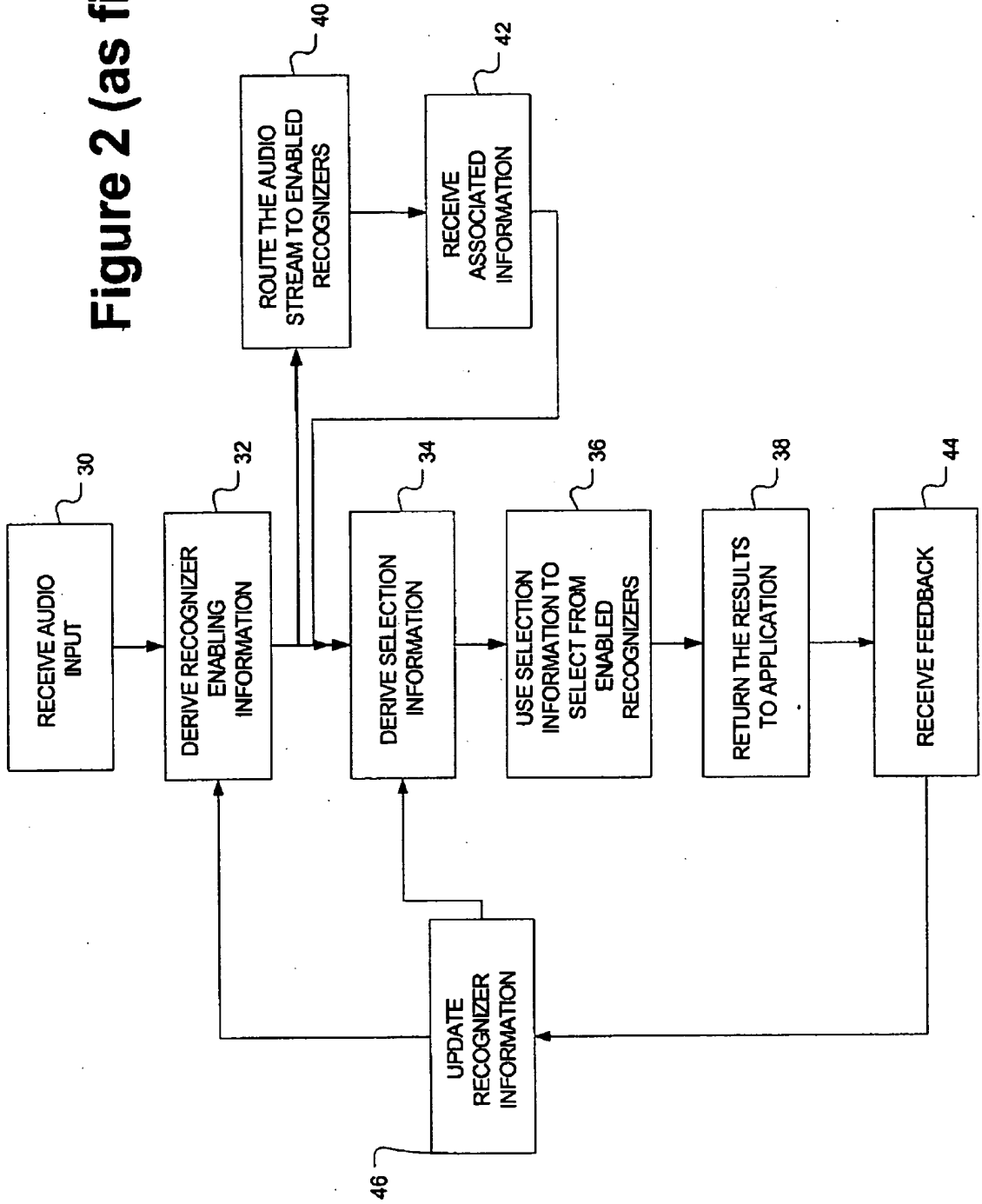
Dated: 3/17/2004



Joni D. Stutman-Horn
Patent Attorney
Intel Corporation
Registration No. 42,173
(703) 633-6845

c/o Blakely, Sokoloff, Taylor &
Zafman, LLP
12400 Wilshire Blvd.
Seventh Floor
Los Angeles, CA 90025-1026

Figure 2 (as filed)



Accepted
G. M.
24 MAY '04